GridFTP

Challenges In Data Transport

John Bresnahan

bresnaha@mcs.anl.gov

Argonne National Laboratory

The University of Chicago

Challenges Past and Future

- Standards

- Throughput

- Robustness

- Extensibility

- Security

- Scalability

Standards

- Interoperability
  - ◆ Big selling point for adoption
- GridFTP 1
  1) Designed
  2) Implemented
  3) Released/Deployed/Used
  4) Standardized
- GridFTP 2
  1) Standardized
  2) Imple....

Throughput

- It had to be fast
  - ◆ GridFTP was sold on speed
  - ◆ Other features eliminate excuses not to use
- Fast varies with the environment
  - ◆ LANs, WANs, Long Fat Pipe
  - ◆ Must be able to configure and exchange protocols
    - TCP window sizes, UDP based protocols
  - ◆ See extensibility

Lost Of Small Files

- 1 large file is easy (but less prevalent)
  - ◆ Overhead to payload ratio is high

- 1 data set partitioned in many little files
  - ◆ Overlap control overhead in data payload
  - ◆ Pipelining
  - ◆ Concurrent sessions
  - ◆ Data channel caching

Robustness

- It has to work ALL the time
  - Hard to get a solid stable code base
    - Harder to extend it
  - Race conditions
- But of course it can't
  - Recover from errors
  - Check point transfers
  - A session crash can't be a service crash
    - Fork()/setuid()/exec()

Extensibility

- Everything has a version 2.0
  - ◆ Even Garbage
- Clean/safe abstractions
  - ◆ ability to add significant features without compromising stability
- In the right place
  - ◆ A balance between control and ease of development.
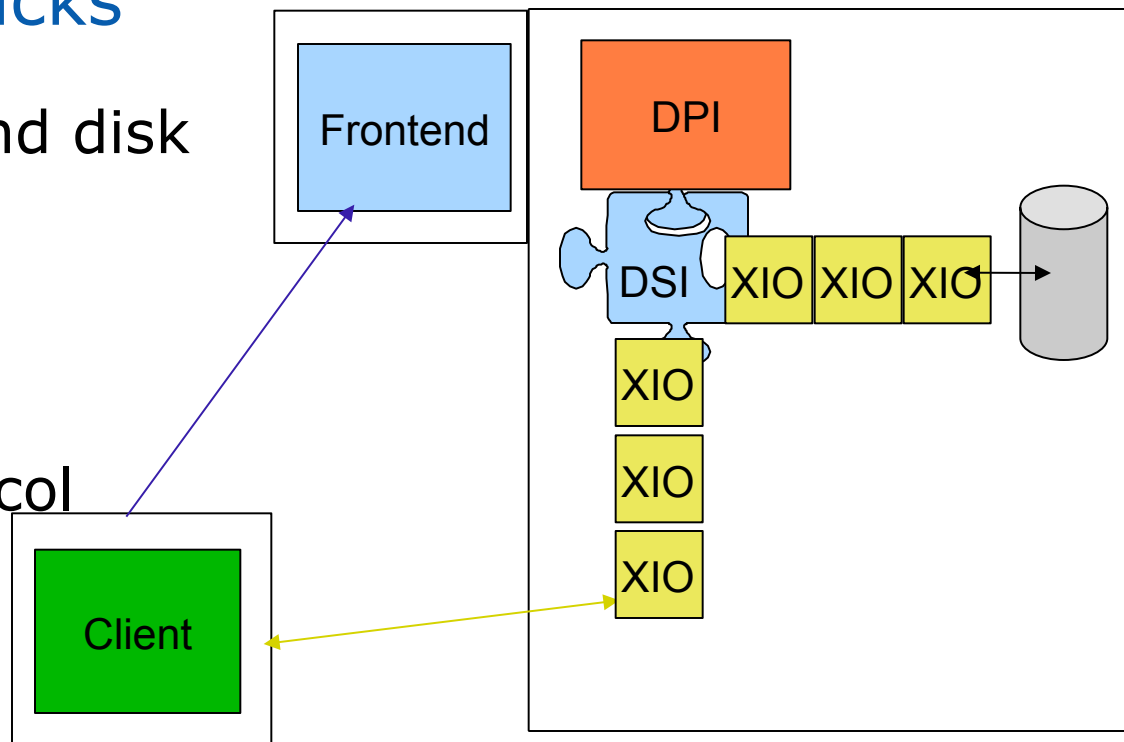 - ◆ XIO
  - ◆ DSI

XIO

- A stack of data interceptors
  - ◆ Filesystem
  - ◆ Data channel
  - ◆ Alter/monitor read/write buffers
- Treats the data as a stream
- Options at open only
- Application treats it as it would a file stream

XIO Driver Stacks

- All data passes through XIO driver stacks
  - to network and disk
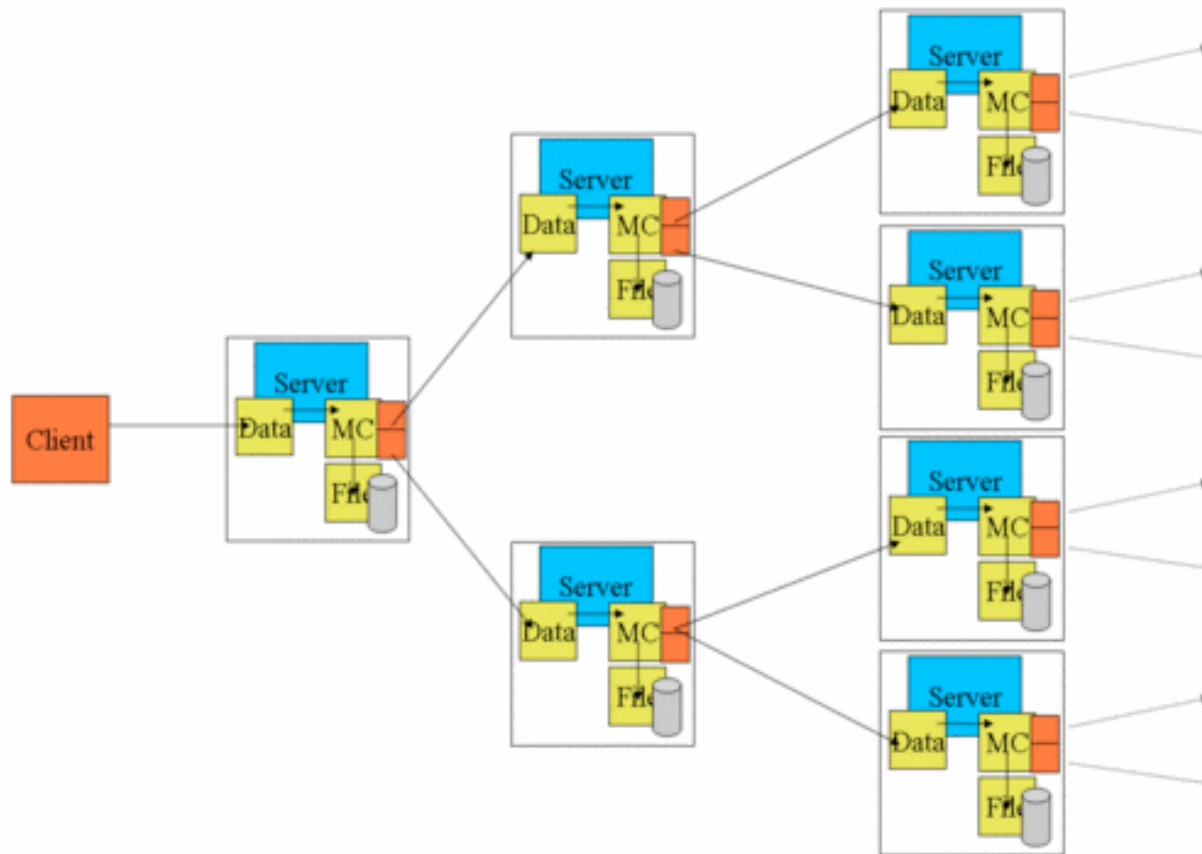 - observe data
  - change data
  - change protocol

GridFTP XIO Extension Examples

- Netloger
  - ◆ Observes and times events for bottleneck detection

- Bandwidth Rate limiter
  - ◆ Throttles the rate buffers are passed along

- Multicast
  - ◆ Forward the buffer to many places

- UDT
  - ◆ Switch out transport protocols

Multicast

- Prototyped in a week

GridFTP over UDT

	Argonne to NZ Throughput in Mbit/s	Argonne to LA Throughput in Mbit/s
Iperf – 1 stream	19.7	74.5
Iperf – 8 streams	40.3	117.0
GridFTP mem TCP – 1 stream	16.4	63.8
GridFTP mem TCP – 8 streams	40.2	112.6
GridFTP disk TCP – 1 stream	16.3	59.6
GridFTP disk TCP – 8 streams	37.4	102.4
GridFTP mem UDT	179.3	396.6
GridFTP disk UDT	178.6	428.3

Data Storage Interface (DSI)

- Intercept all file system calls

  - stat, remove, mkdir, send, receive, …

- Must handle the I/O for the FS

  - Harder to write

  - Much more flexibility

- Examples
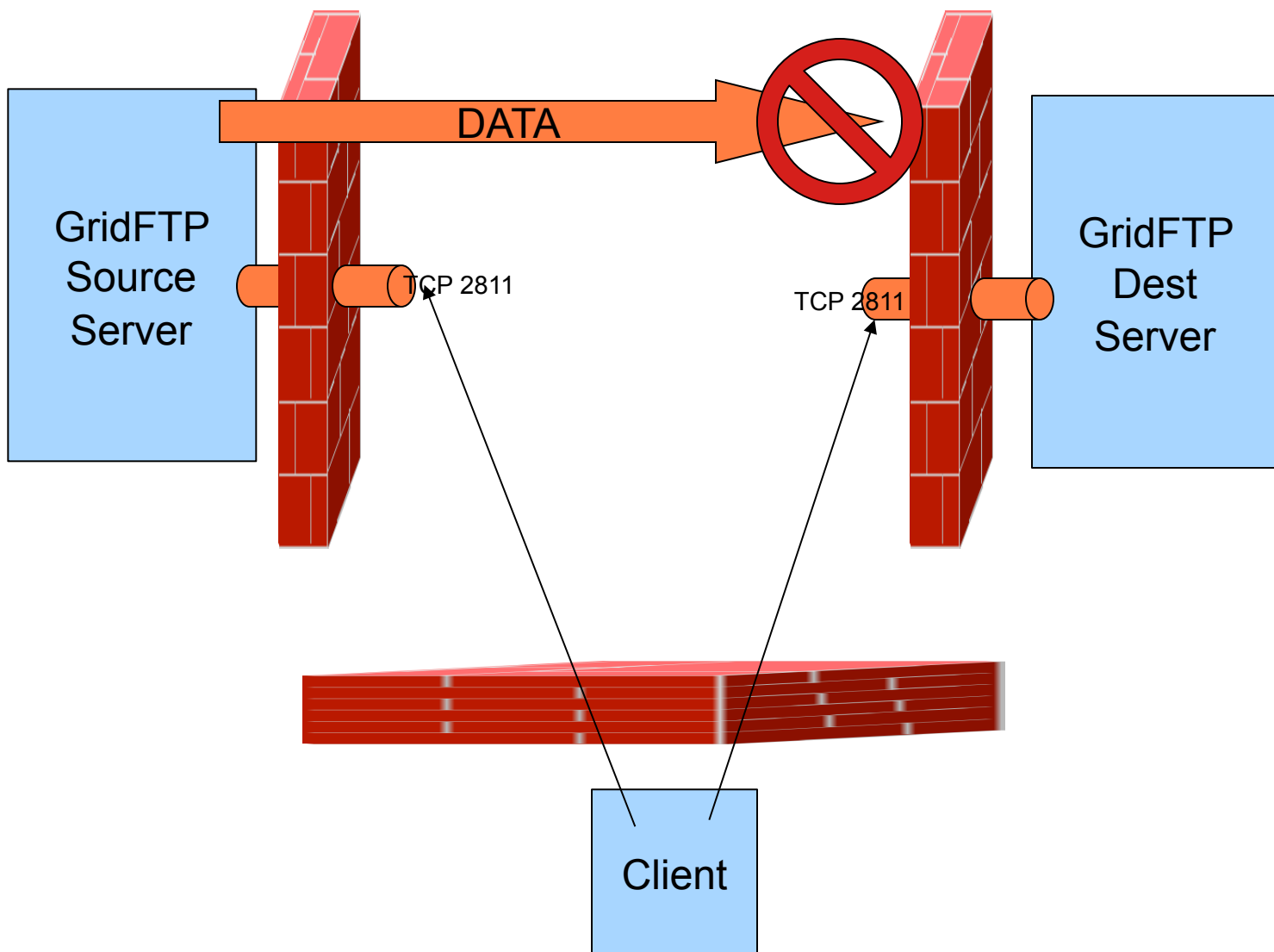
  - HPSS, SRB, proxy/striping

Security

- Protection vs. Ease of use
  - ◆ GSI and CAs were hard for many users
- Speed vs. protection
  - ◆ Users area happy with a minimal amount of data channel protection
- Warm fuzzies
  - ◆ Simple and unsafe mode
- Flexibilty
  - ◆ XIO drivers handle security
    - Still hard to extend
  - ◆ GridFTP over SSH
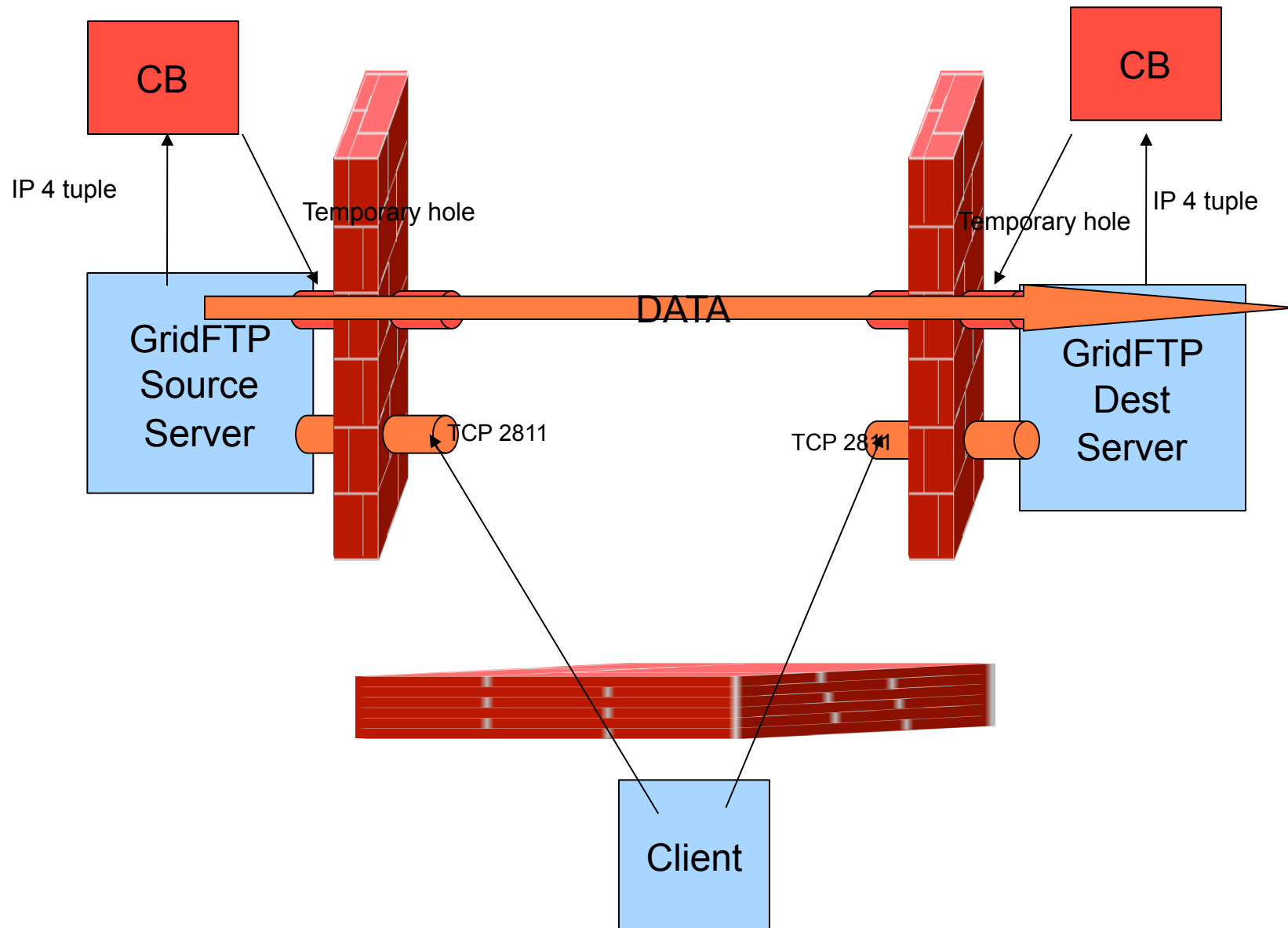   - A big win for many users

Firewalls

- Punching through
  - ◆ Control channel is statically assigned
  - ◆ Data channels dynamically assigned
- 1 way firewall (and NAT)
  - ◆ Automatic traversal
  - ◆ Simultaneous Open/TCP splicing
  - ◆ STUN
- 2 way firewall
  - ◆ Use a broker to create a route
  - ◆ Negotiate the local ports
    - new protocol needed
  - ◆ Hooks in GridFTP to contact a broker at the right time

Outgoing allowed

the globus alliance
www.globus.org

GridFTP Source Server

GridFTP Dest Server

DATA

TCP 2811

TCP 2811

Client

Connection Broker

the globus alliance
www.globus.org

CB

CB

IP 4 tuple

IP 4 tuple

Temporary hole

Temporary hole

DATA

GridFTP
Source
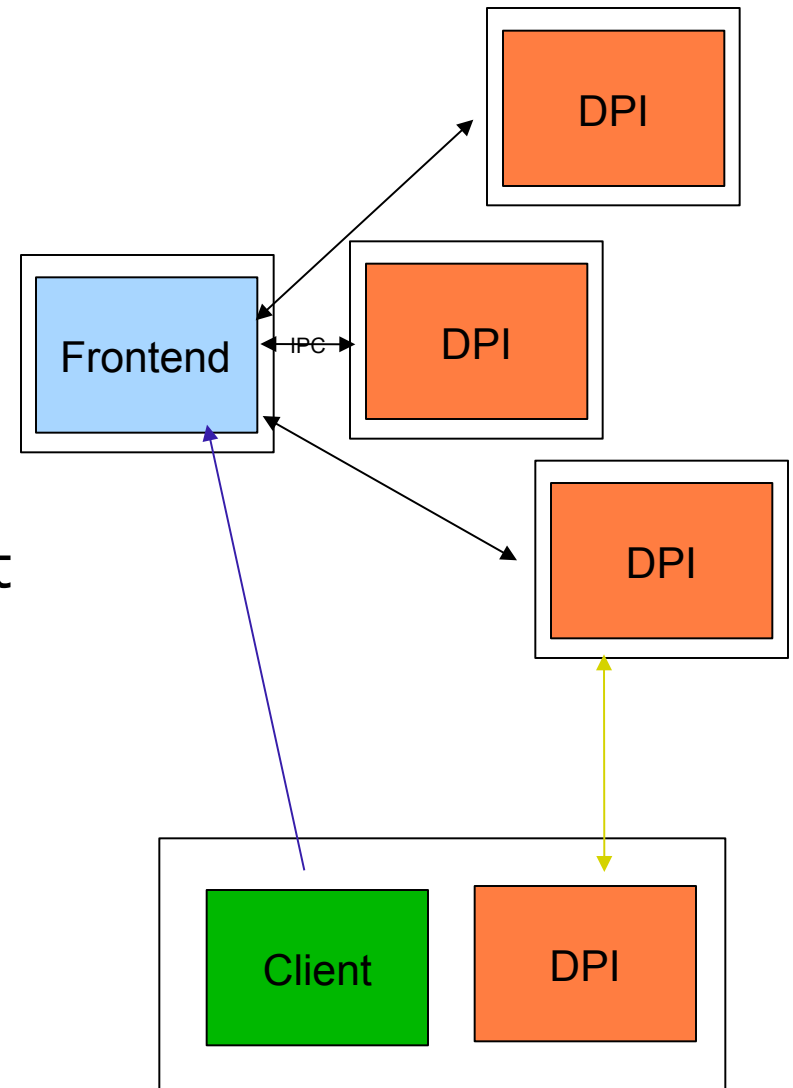Server

GridFTP
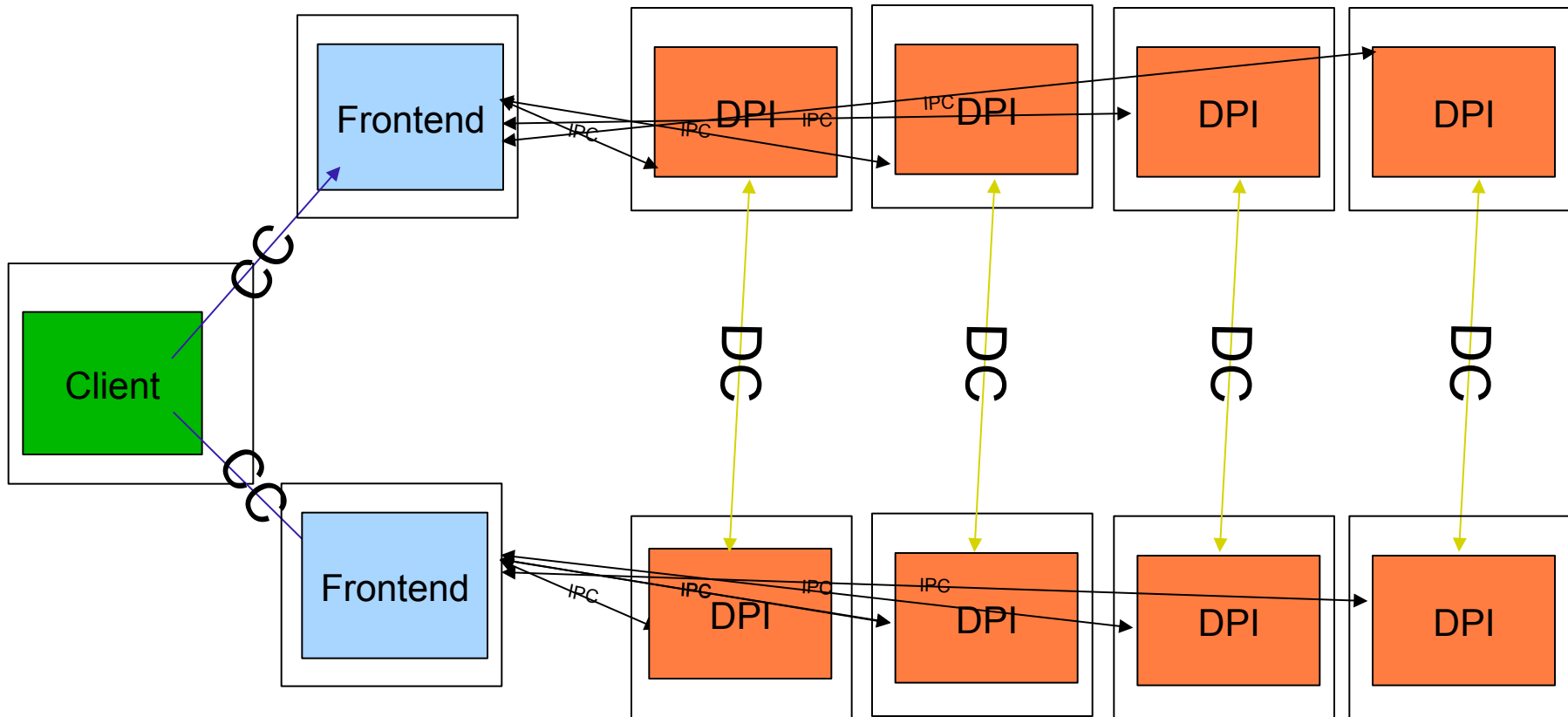Dest
Server

TCP 2811

TCP 2811

Client

Scalabilty

- Striping
  - Multi-host coordinated transfers
  - You give us the hardware, we'll give you the bandwidth
- Load balancing proxy
- Dynamic backends

Proxy Server

- The separation of processes buys the ability to proxy

  - Allows for load balancing

  - Frontend can choose from a pool of DPIs to service a client request

Striping

Multi-core

- CPU/to NIC ration increasing
  - ◆ Treat each core as a stripe
  - ◆ Parallel stream on each core
- Fully encrypted transfers at network speeds
  - ◆ all the security, none of the perf loss
- Compression
  - ◆ Faster than network speed transfers

Conclusion

- Past success
  - Robustness
  - Throughput
  - Standard
- Future (+=)
  - Scalable
  - Secure
  - Extensible
- http://www.gridftp.org
- bresnaha@mcs.anl.gov